

Counting the number of planets around GJ 581

False positive rate of Bayesian signal detection methods

Mikko Tuomi^{★1,2} and James S. Jenkins³

¹ University of Hertfordshire, Centre for Astrophysics Research, Science and Technology Research Institute, College Lane, AL10 9AB, Hatfield, UK

² University of Turku, Tuorla Observatory, Department of Physics and Astronomy, Väisäläntie 20, FI-21500, Piikkiö, Finland

³ Departamento de Astronomía, Universidad de Chile, Camino del Observatorio 1515, Las Condes, Santiago, Chile

Received XX.XX.2012 / Accepted XX.XX.2012

ABSTRACT

Context. The four-planet system around GJ 581 has received attention because it has been claimed that there are possibly two additional low-mass companions as well – one of them being a planet in the middle of the stellar habitable zone.

Aims. We re-analyse the available HARPS and HIRES Doppler data in an attempt to determine the false positive rate of our Bayesian data analysis techniques and to count the number of Keplerian signals in the GJ 581 data.

Methods. We apply the common Lomb-Scargle periodograms and posterior sampling techniques in the Bayesian framework to estimate the number of signals in the radial velocities. We also analyse the HARPS velocities sequentially after each full observing period to compare the sensitivities and false positive rates of the two signal detection techniques. By relaxing the assumption that the radial velocity noise is white, we also demonstrate the consequences that noise correlations have on the obtained results and the significances of the signals.

Results. According to our analyses, the number of Keplerian signals favoured by the publicly available HARPS and HIRES radial velocity data of GJ 581 is four. This result relies on the sensitivity of the Bayesian statistical analysis techniques but also depends on the assumed noise model. We also show that the radial velocity noise is actually not white and that this feature has to be accounted for when analysing radial velocities in a search for low-amplitude signals corresponding to low-mass planets.

Conclusions. Using the Bayesian analysis techniques, we did not obtain any false positives generated by noise when analysing subsets of the HARPS velocities. This suggests that any signals we detect using the Bayesian detection criteria are genuine signals present in the data. However, some such signals could be misinterpretations caused by an insufficiently accurate noise model or by (systematic) noise when independent confirmation of the signals with other data sources is not possible.

Key words. Methods: Statistical, Numerical – Techniques: Radial velocities – Stars: Individual: GJ 581

1. Introduction

The nature of the planetary system around GJ 581 has been discussed in the literature by several authors and the preferred orbital properties and number of planets in the system have been revised whenever new data has become available. While such accumulation of knowledge is the very essence of science, contradicting results, especially those based on robust statistical arguments, have seen the system around GJ 581 become a planetary system with four established planets in its orbits and another two hypothetical ones that have been shown to exist in some studies but have been shown detections of false positives in others.

Astronomers first started counting the planets orbiting GJ 581 in 2005 when Bonfils et al. (2005) reported they had found a Neptune-mass planet ($m_p \sin i = 16.6 M_\oplus$) with an orbital period of roughly 5.4 days. This discovery, made using the High Accuracy Radial Velocity Planet Searcher (HARPS) spectrograph (Mayor et al., 2003) mounted on the 3.6 m telescope at the European Southern Observatory (ESO) at La Silla, Chile, was soon accompanied by discoveries of two additional super-Earth-mass planets made by the same team in 2007 (Udry et al., 2007). These two additional planets were reported to have min-

imum masses of 5.0 and 7.7 M_\oplus and orbital periods of 12.9 and 84 days. Udry et al. (2007) argued that the outer planet (GJ 581 d) was on a slightly eccentric orbit ($e = 0.20 \pm 0.10$) but also noted that its orbit could be consistent with a circular one as was the case for the other two planets.

In 2009, Mayor et al. (2009) reported that additional HARPS velocities revealed the existence of a fourth planet orbiting GJ 581 with an orbital period of 3.1 days. Mayor et al. (2009) detected this planetary companion from 119 velocity measurements and claimed it to have a low minimum mass of 1.9 M_\oplus . They also revised the orbital period of GJ 581 d to 66.4 days and stated that the period of 84 days obtained by Udry et al. (2007) was in fact a yearly alias of the 66.4-day signal. This is a textbook example of problems aliasing can cause to detections of periodic signals in radial velocities and other time series.

Counting the planets orbiting GJ 581 got complicated in 2010 and 2011 when Vogt et al. (2010) claimed that a combined dataset with the 119 HARPS measurements of Mayor et al. (2009) and another set of 122 High Resolution Echelle Spectrograph (HIRES; Vogt et al., 1994) velocities from the 10 m Keck-II telescope at Hawaii actually contained the periodic signatures of – not only the four planets observed by Mayor et al. (2009) – but also two other signals corresponding to low-mass planets with minimum masses of 7.0 and 3.1 M_\oplus and

[★] The corresponding author, e-mail: mikko.tuomi@utu.fi; m.tuomi@herts.ac.uk

orbital periods of 433 and 36.6 days, respectively. The immediate conclusion of Vogt et al. (2010) was that the latter might be a good candidate for a habitable planet with its Earth-like mass and its orbit in the middle of the stellar habitable zone of GJ 581. However, the analyses of Tuomi (2011) and Gregory (2011) soon revealed that the existence of this candidate habitable world was not supported by the data and that the evidence in favour of the outer planet with an orbital period of 433 days was at most suggestive. Gregory (2011) and Tuomi (2011) concluded that there were five and four planets, respectively, in the system depending on the subjective choices of the probabilistic detection threshold and the prior densities of the parameters of the statistical models. Yet, the existence of the four planets detected by Mayor et al. (2009) was easily verified using the Bayesian statistical tests of Tuomi (2011) and Gregory (2011).

Other authors have also re-analysed the HARPS and/or HIRES data of Mayor et al. (2009) and Vogt et al. (2010) and received evidence that supports the existence of the hypothetical habitable zone planet GJ 581 g (e.g. Anglada-Escudé & Dawson, 2011; Tadeu dos Santos et al., 2012). However, it has been clear from the start, as also stated by almost every author discussing the system and estimating the number of planets in it, that only additional data can firmly verify or falsify the existence of the two proposed companions GJ 581 f and g.

Such data was made available by Forveille et al. (2011) and increased the amount of HARPS data of GJ 581 by another 121 velocities. This increase made the total number of HARPS precision velocities 240. However, despite attempts of recovering the proposed planets f and g from this data using standard Lomb-Scargle periodograms (Lomb, 1976; Scargle, 1982) and χ^2 minimisations, Forveille et al. (2011) concluded that these two planets could not exist orbiting the star with the masses and orbital properties reported by (Vogt et al., 2010). Recently, a re-analysis of these HARPS velocities backed up by dynamical arguments was found to support the existence of a habitable-zone super-Earth with an orbital period of 32 days and a minimum mass of $2.2 M_{\oplus}$ (Vogt et al., 2012) but another study taking into account the noise correlations did not find evidence in favour of the existence of this signal in the same data and also casted doubt on the significance of the signal corresponding to GJ 581 d (Baluev, 2012). Therefore, considering the different estimates for the number of planets (parameter k) favoured by the data in the very recent history suggests that this debate is far from over.

If of planetary origin, any periodic variations in radial velocities should be present in data obtained using different telescope-instrument combinations. The availability of such independent confirmation is essential to exclude spurious signals that can be caused by instrument instability and insufficient noise modelling of the features in the measurement noise of the instrument.

Using the combined set of HARPS and HIRES velocities of GJ 581, we perform a global search for low-amplitude signals in an attempt to see whether the conclusions of Forveille et al. (2011), stating that $k = 4$, hold for results obtained using Bayesian analysis techniques (e.g. Tuomi, 2011, 2012; Tuomi et al., 2011; Tuomi & Jones, 2012). We also re-analyse the 240 HARPS velocities of Forveille et al. (2011) to (1) enable comparison of our results with those of Bonfils et al. (2005), Udry et al. (2007), Mayor et al. (2009), Forveille et al. (2011), Vogt et al. (2012), and Baluev (2011); to (2) quantify the sensitivity of the Bayesian detection methods w.r.t. the frequentist periodogram tools used throughout the literature; and to (3) see whether the apparent sensitivity of the Bayesian detection thresholds is prone to false positives in practice. Our goal is to re-

ceive a reliable estimate for k in the case of GJ 581 radial velocities and to determine the significances and properties of possible additional signals if $k > 4$. In our Bayesian analyses, we use the “standard” white noise models assuming independent and identically distributed measurements with Gaussian white noise but we also test models with red noise features, i.e. we model the data using a moving average (MA) model that accounts for the noise correlations in the data in a time-scale of few dozen days (e.g. Baluev, 2012; Tuomi et al., 2012a,b).

2. Statistical analysis

2.1. Bayesian statistics

When assessing the optimal value of k , i.e. the number of signals statistically significantly present in a time series, we take advantage of Bayesian model selection that enables the simultaneous comparison of any number of models – in this particular case, models with different k . We denote these models as \mathcal{M}_k , for all $k = 0, \dots, p$, and calculate the corresponding probabilities of each model in a standard way as

$$P(\mathcal{M}_k|m) = \frac{P(m|\mathcal{M}_k)P(\mathcal{M}_k)}{\sum_{i=0}^p P(m|\mathcal{M}_i)P(\mathcal{M}_i)}, \quad (1)$$

where m denotes the measurements; $P(m|\mathcal{M}_i)$ is the Bayesian evidence, or probability of obtaining the data if the model \mathcal{M}_i indeed was a “correct” model; and $P(\mathcal{M}_i)$ is the prior probability of the model \mathcal{M}_i that quantifies the subjective level of confidence we have on the validity of this model *a priori* to analysing the measurements. As can be seen in Eq. (1), the probabilities are scaled in such a way that the sum of the individual probabilities of models \mathcal{M}_k , for $k = 0, \dots, p$, is unity indicating that any posterior probabilities obtained using this equation are dependent on the selected set of candidate models and only reflect the *relative* probabilities of the models in this set.

Finding the optimal value of k based on model selections using Eq. (1), however, is not as such a sufficient way of counting the periodic signals in a given data set. While a given k might receive the greatest posterior probability, the claim that there really are k signals in the data should not be made lightly. This is especially the case if the probability of a model with some lower value of k receives estimates of similar magnitude. For this reason, we require that the probability of a model \mathcal{M}_{k+1} has to be at least 150 times that of a model \mathcal{M}_k to conclude that there is evidence in favour of the existence of $k + 1$ Keplerian signals in the data. This threshold corresponds to an interpretation of “strong evidence” given by Kass & Raftery (1995) to make decisions based on data. While completely subjective, we adopt this threshold but do not consider signals to exist in the data unless the other two criteria of Tuomi (2012) are also satisfied, i.e. that the estimates of the velocity amplitude parameter K are statistically significantly greater than zero and that the period P is constrained from above and below. If all these criteria are satisfied, we state that the existence of a signal is supported by data.

While the prior probabilities in Eq. (1) are only simple numbers, the Bayesian evidences are in fact complicated integrals over the whole parameter space of model parameters that depend on the value of the product of a likelihood function and a prior probability density of the model parameters, i.e. the formulation of the statistical model and the physical interpretation of its parameters given the available data. For a given model, this integral, sometimes called the marginal integral, is of essence in any model comparison problems in the Bayesian context and the

ability to compare models in a probabilistic way depends on the ability to estimate this integral. We write this integral as

$$P(m|\mathcal{M}_k) = \int_{\theta_k \in \Omega_k} l(m|\theta_k, \mathcal{M}_k) \pi(\theta_k|\mathcal{M}_k) d\theta_k, \quad (2)$$

where the parameters of the k th model, θ_k , in the parameter space Ω_k , are interpreted using the likelihood function l and their prior probability density π that describe the mathematical details as well as the physical interpretation of the statistical model \mathcal{M}_k . This is also the reason we write the likelihood and the prior as conditioned on the model.

We estimated the integral in Eq. (2) using the truncated posterior mixture (TPM) estimate of Tuomi & Jones (2012). This estimation requires the availability of a statistically representative sample from the parameter posterior density and we drew such samples using the adaptive Metropolis algorithm (Haario et al., 2001) that is simply an adaptive version of the famous Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm (Metropolis et al., 1953; Hastings, 1970). This sampling technique has been used successfully in e.g. Tuomi (2012).

Apart from the chosen set of candidate models (i.e. the definition of their likelihood functions), the only other subjective issue in Bayesian statistics is the prior information quantified using prior probability densities of the model parameters and prior probabilities of the different models. Standard (frequentist) analyses of radial velocity data actually incorporate prior information as well, which makes them somewhat Bayesian in reality, but do not easily enable testing different prior opinions in practice. For instance, when searching for a solution using typical χ^2 minimisations, the amount of excess noise in the measurements, or parameter σ_J , is commonly fixed to an *a priori* selected value (e.g. σ_0), which corresponds to setting the prior probability of this parameter equal to a delta-function as $\pi(\sigma_J) = \delta(\sigma_J - \sigma_0)$. This prior assumes that only a value of $\sigma_J = \sigma_0$ is a possible one and ignores the rather likely possibility that this parameter might not be equal to σ_0 . If this prior choice was made in the Bayesian context, any scientifically oriented reader would immediately dismiss this choice as too limiting and would not consider the obtained results trustworthy. However, such *prior limitations*, i.e. setting prior densities to zero in chosen subsets of the parameter space, are commonly used in statistical analyses but should be used with care and only when the physical interpretation of the parameters justifies the limitations.

Similar “hidden” prior assumptions are made throughout the literature announcing detections of exoplanets around nearby stars. For instance, finding a solution to some data by assuming that planetary eccentricities are equal to zero represents a similar prior choice. It could be argued that in “dynamically packed” systems only eccentricities close to zero are viable by enabling the long-term stabilities of these systems, but the possibility that some non-zero orbital eccentricities could also provide stable systems cannot be ignored because enabling eccentricities as free parameters effectively increases the number of free parameter in the model by $2k$ compared to a model with fixed eccentricities and this can have considerable effects on the significances of the obtained solutions. As fully Bayesian, these hypotheses could be easily tested by treating the model with eccentricities fixed to zero as another model in the set of candidate models. In practice, this can be seen as a comparison of different prior densities of the model parameters.

The argument of hidden prior information can in fact be taken even further. Any non-Bayesian analyses correspond to Bayesian ones with the assumption that the prior probability densities of model parameters are actually “flat”, i.e. uniform

distributions in the parameter space. This is a typical uninformative choice for a prior density and is applied widely in the statistical literature to various parameter estimation problems. However, it still *is* a subjective choice and it *does* have an effect on the obtained results. For instance, a typical non-linear transformation of the parameter vector from θ to θ' reveals the consequences of this prior choice. If the prior for parameter θ is a uniform density, that for θ' has to be chosen by using the Jacobian of the corresponding transformation to receive the same results. If the prior of θ' is chosen uniform as well, the obtained analysis results would not be the same for these two parameterisations but would depend heavily on the chosen transformation, which emphasises the fact that prior densities are important entities that cannot be ignored when performing statistical analyses. This in turn implies that the only logically consistent framework of statistics is the Bayesian one. While these issues are discussed widely in the statistical literature, we refer to the excellent book *Statistical Decision Theory and Bayesian Analysis* by J. O. Berger (Berger, 1980).

Because priors are subjective, we present our results only based on the priors we choose. In practice, we adopt the same prior probability densities of model parameters that were used in Tuomi (2012). We also use the same prior model probabilities as Tuomi (2012) to take into account the fact that we believe finding $k+1$ planets in any given system is less likely than finding k of them. Because posterior samplings can be used to obtain estimates of the parameter probability densities that enables the computation of any point- and uncertainty estimates, we report all the results using the maximum *a posteriori* (MAP) estimates and the corresponding 99% Bayesian credibility sets (BCSs; as defined in e.g. Tuomi & Kotiranta, 2009).

2.2. Periodograms

Because it is useful to calculate the periodograms of time series, especially when searching for periodic signals, we discuss the resulting power spectra as well for the GJ 581 velocities. Identifying the strongest powers in the periodogram, such as the famous Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982), helps choosing the initial states of posterior samplings by pointing towards the regions in the period space where periodic signals are the most likely to be found. The ability to use information from periodogram analyses can help decreasing the burn-in period, i.e. the length of the Markov chain before it has converged to the global solution, of MCMC samplings considerably. However, because periodogram analyses are commonly used to assess the number k in practice, we first discuss the caveats of such assessments.

As noted by Tuomi (2012), periodograms and related analysis methods rely on the assumption that the model residuals have been calculated correctly. For instance, when searching for a $k+1$ th signal in a noisy time series, the residuals are commonly calculated by assuming that there are in fact only k periodicities in the data. This obvious contradiction of the assumption and the statistical test performed to search for another signal is not only made throughout the literature when searching for planets using the radial velocity technique, but its effects on the detectability of periodic signals have not received as much attention in the same literature as they should. A clear example is the case of HD 10180 that was reported to host six, possibly seven, planets orbiting it by Lovis et al. (2011). They used power spectra in their attempt of assessing the number k for this star and fell a victim of the above contradiction of the statistical test and the assumption this test is based on. Contrary, the Bayesian analysis

of the same exact data performed by Tuomi (2012) concluded that the preferred estimate for k is actually as high as nine. This example only demonstrates the weakness of the periodogram-based methods in assessing k but can also be seen as an obvious consequence of the assumption: if it is assumed that there are k signals in a data set, the statistical tests performed based on this assumption will certainly be biased in favour of this assumption being true and prevent the detection of a $k + 1$ th signal unless it is a strong one and clearly present in the data.

In an attempt to overcome the above shortcomings of periodogram analyses, Anglada-Escudé & Tuomi (2012) proposed a generalised version of the periodogram defined by Cumming (2004) for the purpose of detecting multiple periodicities in time series. This *recursive periodogram* (Anglada-Escudé & Tuomi, 2012) can be used to obtain the power spectrum for model residuals by simultaneously adjusting the parameters of the previously detected k signals for each test frequency of the $k + 1$ th one. Clearly this method is not prone to such a severe contradiction as the traditional periodograms but it is still not free of its own conceptual problems such as the fact that one has to move sequentially by increasing k one step at a time when searching for several periodicities. While certainly an improvement to the classical periodograms, the consequences of hidden prior information cannot be addressed using this modified periodogram either.

We report the results of standard Lomb-Scargle periodogram analyses together with those from the Bayesian analyses. We also report the resulting false alarm probabilities (FAPs) of the signals we detect in the GJ 581 velocities obtained using the HARPS spectrograph.

3. Sequential analysis of HARPS velocities

The HARPS velocities provide an interesting opportunity to compare the sensitivity of the Bayesian analysis methods to the traditional periodogram ones used throughout the literature. For this reason, we construct a timeline for GJ 581 where we analyse the available HARPS data after each yearly observing period. We also analyse the data sets of Bonfils et al. (2005), Udry et al. (2007), and Mayor et al. (2009) and compare the best estimate of k with an estimate that *could* have been received had Bayesian tools been applied to these velocities since 2005.

We show the number of HARPS velocities after each observing period and at the time when the three studies describing detections of the four planets orbiting GJ 581 and reporting the data were published (these three studies are: Bonfils et al., 2005; Udry et al., 2007; Mayor & Queloz, 1995). The numbers of measurements and data baselines of these subsets of HARPS data are shown in Table 1. In the following subsections we describe briefly the results of our analyses of the subsets of HARPS data in Table 1. We note, that we assign numbers to the signals we detect by calling the shortest periodicity as signal 1, the second shortest as signal 2 and so forth. We also refer to the signals corresponding to the four planet candidates by using the letters b, c, d, and e as in Mayor et al. (2009).

3.1. Bonfils et al. data

When analysing the Bonfils et al. (2005) data set, our posterior samplings could rapidly verify that it contains the clear planetary signature of GJ 581 b. This signal was seen easily with posterior samplings (with $P(k = 1) \approx 5.9 \times 10^{10} P(k = 0)$) and periodograms (FAP < 0.0001) and Bonfils et al. (2005) were also

Table 1. Number of signals detected from the subsets of HARPS data as a function of time as reported in the literature (k_1), using Lomb-Scargle periodograms (k_2), and using Bayesian tools (k_3). The corresponding data baselines (T_{obs}) and numbers of data (N) are also shown. The subsets correspond to the published data sets and subsets of Forveille et al. (2011) data set after each full observing period (FOP).

Data subset	N	T_{obs} [days]	k_1	k_2	k_3
Bonfils et al. (2005)	20	440	1	1	1
First FOP	25	457	–	2	2
Second FOP	43	827	–	2	2
Udry et al. (2007)	50	1050	3	3	3
Third FOP	76	1197	–	4	4
Mayor et al. (2009)	119	1570	4	4	4
Fifth FOP	131	1904	–	4	4
Sixth FOP	197	2312	–	4	5
Forveille et al. (2011)	240	2543	4	4	5

very confident about its existence based on their analyses that we assume¹ relied on periodograms and χ^2 minimisations.

Our samplings of a two-Keplerian model also spotted hints of a second periodic signal between roughly 10-16 days. The two-Keplerian model received a significantly greater posterior probability (by a factor of 640) than the one-Keplerian one, though the samplings could not constrain the period of the second signal very accurately. Also, parameter K of this second signal did not differ statistically from zero, which forced us to conclude that there is only one signal significantly present in the Bonfils et al. (2005) data set. Yet, this signal exists with little doubt because taking it into account increases the model probability by a factor of 1.1×10^{11} compared to the model with $k = 0$ despite the low number of measurements (see Table 1).

3.2. First full observing period

The end of the first full observing period provided only five more spectra and corresponding velocities of GJ 581 than the data of (Bonfils et al., 2005). However, this changed the situation with respect to the second signal considerably. The two-Keplerian model became 4.6×10^4 times more probable than the one-Keplerian one, and parameter K of the second signal was significantly greater than zero with $K_2 = 3.6$ [1.5, 5.7] ms^{-1} . Despite the fact that the period of this signal had a density with several modes between roughly 12 and 16 days, it was well constrained from above and below and thus satisfied all the detection criteria. The second periodic signal was also detected by the periodograms and the corresponding FAP of this signal was 0.0075.

The probability densities of the velocity amplitude and period of the second signal are plotted in Fig. 1. The multimodality of the period is clear from the top panel in Fig. 1 but the period is still constrained from above and below and its 99% BCS is [11.98, 16.27] days. We note that the MAP estimate of the period is 12.9 days.

We sampled a three-Keplerian model as well, but unsurprisingly a third periodic signal did not get constrained by the available 25 velocities.

¹ In fact, Bonfils et al. (2005) do not mention the statistical methods they used to obtain their solution.

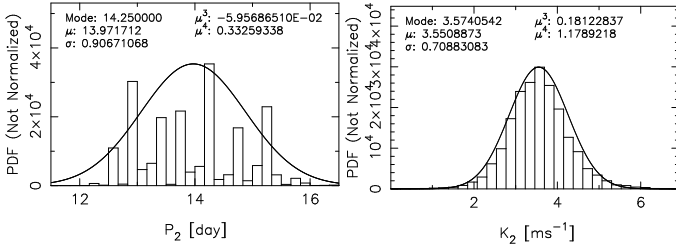


Fig. 1. Distributions of the velocity amplitude and period of the second signal (corresponding to GJ 581 c) as obtained from the partial HARPS data after the first full observing period. The solid curve shows a Gaussian function with the same mean and variance.

3.3. Second full observing period

At the end of the second full observing period, the amount of velocities had increased to 43, and we could easily sample models with $k = 0, \dots, 3$. For this data set, the three-Keplerian model implied that there are at least three periodicities in the data but the period of the third signal turned out to have a density with three modes. These modes were found at periods of 42.3, 67, and 84 days. We could not set constraints to the period, and could not be sure whether our Markov chains had indeed converged to this multimodal density – but assuming they had, the three-Keplerian model would have had a probability of 620 times that of the two-Keplerian one whose parameter space was well sampled and provided two clear periodicities at 5.37 and 12.9 days.

When we sampled the parameter space of a four-Keplerian model, however, we could identify a solution for a fourth period at 3.1 days. This solution, while only suggestive because the period was not constrained from above or below but was only apparent as a global maximum in the otherwise “flat” probability density, actually helped constraining the third periodic signal from above and below according to our detection criteria. While this subset of HARPS data could not be used to claim a detection of four periodicities, the results suggested that only a modest increase in the number of available data would be likely to provide detections of four signals. Yet, according to our detection criteria, we could spot only two signals confidently in this data set, namely, those with periods of 5.37 and 12.9 days. These two signals were also detected by the periodogram analyses (Table 1).

3.4. Udry et al. data

Udry et al. (2007) reported that the Bonfils et al. (2005) data already actually suggested there might be an additional companion orbiting GJ 581 with an orbital period of 13 days with a modest significance. With additional 30 data points and improved barycentric corrections, their periodogram analyses and χ^2 minimisations pointed towards two new periodic signals, corresponding to two new planets orbiting GJ 581 with orbital periods of 13 and 84 days.

These two signals were indeed significantly present in the Udry et al. (2007) data according to our Bayesian analyses. However, we could go further than that and identify a fourth periodicity at 3.14 days corresponding to the fourth planet candidate reported by Mayor et al. (2009). We show the probability densities of velocity amplitude and period of the fourth signal (Fig. 2) to demonstrate that the fourth periodicity was detected as a clear probability maximum but it did not quite satisfy the detection criteria, namely, parameter K of the 3.14 day signal was

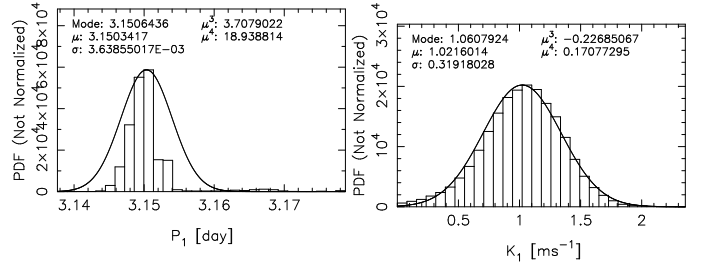


Fig. 2. As in Fig. 1 but for the velocity amplitude and period of the signal with the shortest periodicity (corresponding to GJ 581 e) from the Udry et al. (2007) data.

not significantly greater than zero by having a tail that extended all the way to zero in the amplitude space (Fig. 2, top panel). However, the signals reported by Udry et al. (2007) were clearly detected in the data.

The periodogram analyses revealed the existence of three signals in this data set as well. The third most significant signal was found at 82.6 days with a FAP of 0.001 but the interpretation of the origin of this signal corresponding to the reported period of 84 days was less clear than claimed by Udry et al. (2007) because the sixth strongest power in the window function was suspiciously close to this period at 87 days making it possible that this periodicity could be due to sampling and not to a genuine signal.

3.5. Third full observing period

After the third full observing period with 76 radial velocities, the 3.14 day signal that was close to satisfying the detection criteria in the Udry et al. (2007) data set was well constrained and satisfied all the detection criteria. With these velocities, the period of the outer planet candidate received a global maximum at 67 days instead of the periodicity at 84 days in the Udry et al. (2007) data set. Essentially, we obtained the same solution that was reported by Mayor et al. (2009) after four full observing periods and 43 more velocities.

To demonstrate that the solution indeed satisfied the detection criteria, we show the posterior densities of the period and velocity amplitude of the “weakest” 3.14 day signal in Fig. 3 (top panels) together with the phase-folded signal (bottom panel). Because the corresponding four-Keplerian solution also had a posterior probability that was 5.4×10^7 times greater than that of the three-Keplerian model, we could conclude that the 76 epochs of HARPS data obtained after three full observing periods were already clearly in favour of a four-Keplerian model. We note that despite the eccentricity prior that penalises high eccentricities, the eccentricity of the outer companion had a MAP estimate of 0.30, yet, its uncertainties were considerable and it was found to be statistically indistinguishable from zero with a 99% BCS of [0, 0.58]. The four signals were also obtained by the periodogram analyses and the FAP of the fourth signal was 0.001, which indicates that this signal was significantly present in the data.

We sampled the five-Keplerian solution and detected two probability maxima for the fifth signal at 25 and 42 days. However, neither of these was found to correspond to a solution that satisfied the detection criteria of another signal – the corresponding periods did not get well constrained and the velocity amplitudes received values were not found distinguishable from zero.

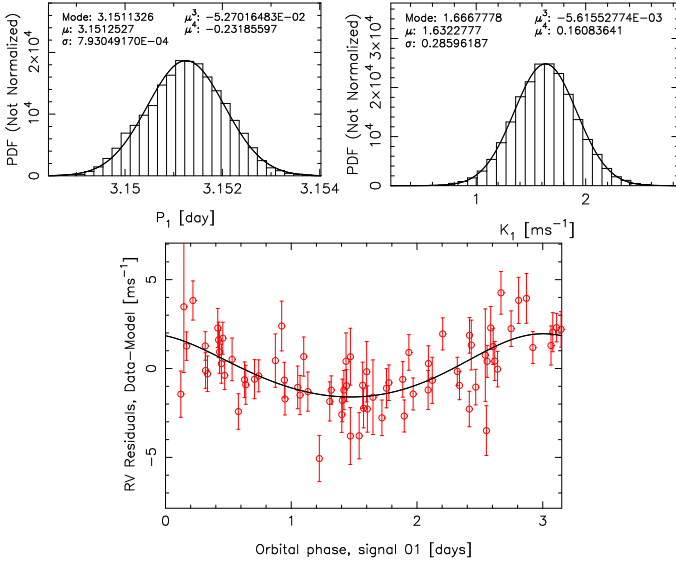


Fig. 3. Probability densities of parameters P and K of GJ 581 e as in Fig. 1 and the phase-folded MAP signal (bottom panel) with the other three signals subtracted.

3.6. Mayor et al. data

Mayor et al. (2009) reported the detection of the fourth and innermost planet candidate orbiting GJ 581 with an orbital period of 3.14 days after having observed 119 spectra of the star. This new data set included velocities based on all the existing spectra after improved spectral reductions and corrections accounting for e.g. a changing internal pressure of the Th-Ar calibration lamps caused by aging. Mayor et al. (2009) stated that they discovered a fourth periodic signal in the velocities using “non-linear minimisations with genetic algorithms” and periodogram analysis.²

We could obtain the same four-Keplerian solution easily using the Mayor et al. (2009) velocities. However, we also tested whether we could spot a fifth signal in this set of 119 radial velocities. Indeed, sampling the parameter space of the five-Keplerian model identified three probability maxima for a fifth periodicity in the data. These maxima were found at periods of 17, 34, and ~ 400 days but the corresponding Markov chains did not actually converge to a posterior density in the sense that their means and variances did not show converge as a function of the chain length. While this might be a result of insufficient chain length, we drew samples of up to few 10^8 members from the posterior density, which should be a sufficient length given the acceptance rate of roughly 10% throughout the samplings (Tuomi, 2012). Also, for the local maxima at 17 and 34 days, the velocity amplitude parameter was not found statistically different from zero, which violates the detection criteria. Similarly, the global maximum of the fifth periodicity at ~ 400 days (likely corresponding to the periodicity reported by Vogt et al. (2010) at 433 days) was found to have an amplitude consistent with zero despite the fact that its MAP estimate was found to be as high as 1.7 ms^{-1} . The signal corresponding to this global maximum was also found to have an extremely poor phase-coverage due to annual gaps in the HARPS velocities.

² Mayor et al. (2009) did not actually specify which measure of model goodness they minimised, which makes it impossible to replicate their results. However, the term “minimisations” that they used suggests that the chosen measure was χ^2 statistics which corresponds to a delta-function prior for the jitter.

3.7. Fifth full observing period

The 131 velocities obtained after the fifth full observing period did not change the solution much with respect to the one obtained using the data set of Mayor et al. (2009). The four signals reported by Mayor et al. (2009) were very clearly present in the data and we moved on to sample the parameter space of a five-Keplerian model. This time, we spotted four different probability maxima for the period of the fifth signal at 32, 42, 190, and 420 days. However, none of these maxima corresponded to solutions satisfying the detection criteria by having either posterior probabilities below the detection threshold or by having velocity amplitudes statistically indistinguishable from zero. Therefore, we could not claim that the data supported the existence of more than four signals according to our detection criteria.

Curiously, the 420 day probability maximum corresponds to the signal reported by Vogt et al. (2010) whereas the 32 day one corresponds to the signal reported by Vogt et al. (2012). These two probability maxima, despite the fact that they do not satisfy the detection criteria, are therefore existing features of the parameter density of the five-Keplerian model. Based on these 131 velocities, however, it is impossible to state whether these probability maxima correspond to actual signals in the data or false positives in the analyses of Vogt et al. (2010) and Vogt et al. (2012).

3.8. Sixth full observing period

After the sixth full observing period, the total amount of HARPS velocities for GJ 581 had increased to 197. These data further improved the constraints of the orbits of the four planets detected by Mayor et al. (2009) and increased their significances w.r.t. the 131 velocities that had been available a year earlier after the fifth full observing period. When sampling the five-Keplerian model, we could identify an interesting solution for the fifth signal.

According to our samplings, there were actually two probability maxima in the five-Keplerian model corresponding to 170 and 190 day periodicities for the fifth signal. These maxima, by being rather close to one another in the period space, enabled drawing a sample from the posterior density and our Markov chains were found to “bounce” between the two maxima yielding a sample from this bimodal solution. This solution satisfied all the detection criteria by having a well constrained period, velocity amplitude that was significantly greater than zero, and a model probability that was 4.7×10^3 times greater than the probability of the four-Keplerian model, which indicates that the favoured value of k is 5 for this data set. While the bimodality is a rather unsatisfactory feature of a probability density – especially when the goal is to quantify the properties of signals corresponding to candidate planets – it is by no means surprising considering that we are searching for periodic signals in the first place. Similar bimodal densities have been observed before when dealing with radial velocities (e.g. Tuomi, 2012). We show the posterior density of the period of the fifth signal in Fig. 4. We could not find any other significant periodicities in this data. Also, we could not detect the fifth signal using the periodogram analyses of the data.

3.9. Forveille et al. data

Forveille et al. (2011) reported that they could spot four and only four periodic signals in their set of 240 HARPS velocities. This was also a simple task for our periodogram analyses as well as for the Bayesian posterior samplings and we continued to sam-

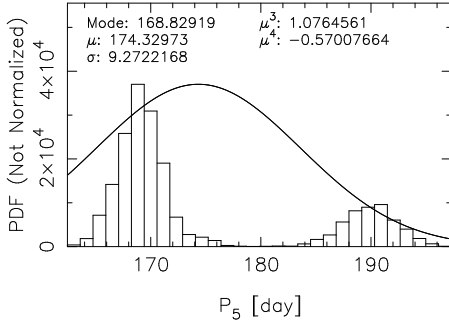


Fig. 4. Probability density of the fifth periodicity in the HARPS data after the sixth full observing period.

ple the parameter space of the five-Keplerian model. These samplings indicated that the 190 day periodicity observed in the data after the sixth observing period gained greater significance while the 170 day period ceased to correspond to a probability maximum. The probability of the five-Keplerian model was found to be 5.7×10^4 times more probable than the four-Keplerian one, which shows that the significance of the fifth periodic signal at 190 days received roughly ten times greater significance than it did for the data after the sixth full observing period. This periodicity was also well constrained and satisfied all our detection criteria. The periodogram analyses could not be used to detect a fifth periodogram in the data confidently.

However, as also reported by Vogt et al. (2012), we also spotted another probability maximum corresponding to a periodicity at 32 days. Treating the 32 and 190 day solutions as separate models, we calculated their respective probabilities. These probabilities were found to be 7% and 93%, respectively, which indicates that the global solution of the five-Keplerian model corresponds to a fifth periodicity at 190 days. Because of the existence of the probability maximum at 32 days, we performed global samplings of a six-Keplerian model but failed to find a signal satisfying the detection criteria. We did, however, spot a probability maximum for the period of the sixth signal at 32 days, but it could not be constrained from above or below, which violates one of our detection criteria. We show the phase-folded fifth signal in Fig. 5 (bottom panel). While the phase-folded signal is not visually very impressive, we also plotted the distributions estimating the period and amplitude of this signal in Fig. 5 (top panels) which indicate that this signal complies with our detection criteria.

4. Analysis of combined radial velocities

In this section we perform a robust analysis of the combined HARPS and HIRES velocities of GJ 581. The baseline of the available 240 HARPS velocities of Forveille et al. (2011) is 2543 days. We plotted the four-Keplerian model residuals of the HARPS velocities in Fig. 6 (top panel). Clearly, these velocities show the remarkable stability (with an RMS of only 1.98 ms^{-1}) of the HARPS spectrograph and indicate that, in addition to obvious yearly gaps, the HARPS data can be expected to provide accurate constraints to possible additional periodic signals in the data.

The residuals of the four-Keplerian model of the HIRES precision velocities of Vogt et al. (2010) are also shown in Fig. 6 (bottom panel). These velocities have an RMS of 2.82 ms^{-1} , which implies an impressive precision, but falls short of the precision of the HARPS velocities for GJ 581. The baseline of the HIRES data is 4000 days, which is almost twice that of the

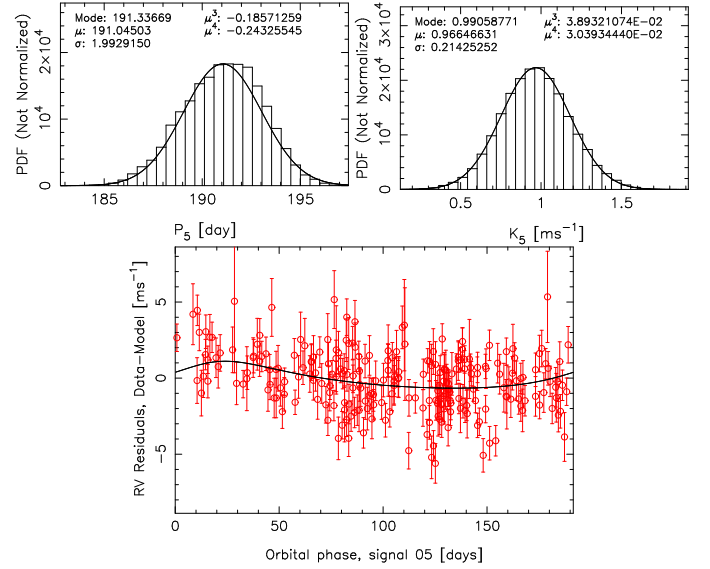


Fig. 5. As in Fig. 3 but for the fifth signal in the Forveille et al. (2011) data.

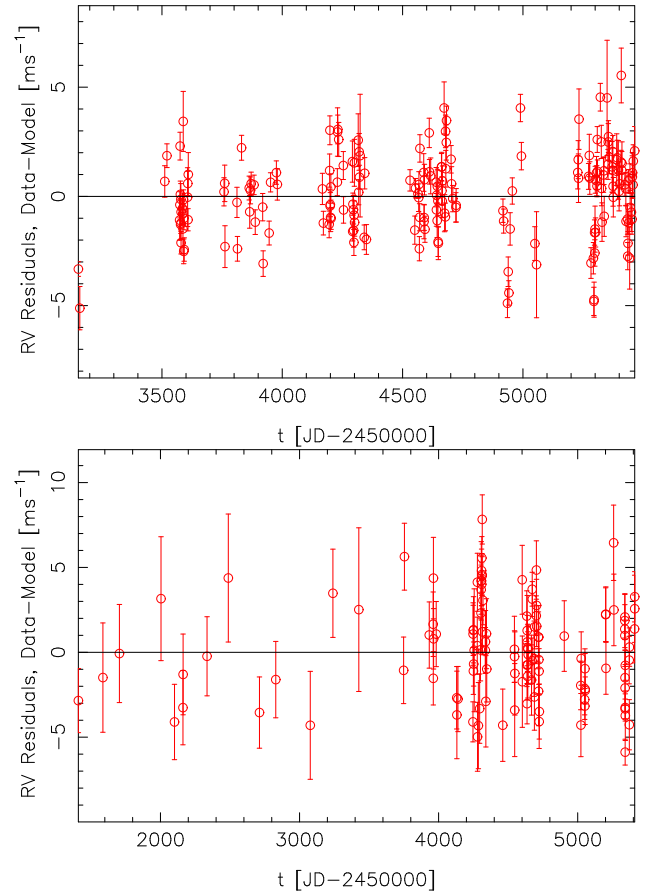


Fig. 6. HARPS (top) and HIRES (bottom) residuals of the four-Keplerian model.

HARPS one. Therefore, combining the HIRES and HARPS velocities would certainly provide greater accuracy than either one of them alone by having a greater overall precision and longer baseline. Also, the differences in the time-distribution of these two sets can certainly help to distinguish genuine Doppler periodicities and their aliases from one another and from spurious

signals generated by noise and insufficient modelling by providing a better sampling than the individual sets.

4.1. White noise model

As we already analysed the HARPS velocities individually in Section 3.9, and because the HIRES velocities have been analysed individually in e.g. Gregory (2011), where it was concluded that they contain two significant periodicities (the two strongest signals at 5.4 and 12.9 days), we move on directly to analyse the combined data set with the pure white noise model.

Using the Gaussian white noise model, the four signals reported by Mayor et al. (2009) were indeed found clearly in the combined HARPS and HIRES velocities and our Markov chains converged to the corresponding four-Keplerian solution with ease. All these four signals were well constrained and satisfied our detection criteria.

When sampling the parameter space of the five-Keplerian model, we searched for signals in the period space between 12.9 and 66.7 days and the space beyond 66.7 days up to a cutoff period of $T_{\max} = 10T_{\text{obs}}$, where T_{obs} is the data baseline, separately and treated them as different models. Effectively, this corresponds to comparing two different prior models, i.e. in this case prior limitations (see Section 2.1), with period ranges as $P \in [12.9, 66.7]$ and $P \in [66.7, T_{\max}]$, where P and T_{\max} have the unit of days, with each other. The cutoff period exceeds the data baseline, but we chose such a long upper limit for the parameter space because signals with periods exceeding the data baseline can be detected in radial velocity data (Tuomi et al., 2009) and this cutoff choice also enables the detection of signals that are only present in the data as trends with no or little curvature. We chose the two other cutoffs of the period space because we expect *a priori* the inner system of GJ 581 to be “dynamically packed” in the sense that because of the existence of planets with orbital periods of 3.1, 5.4, and 12.9 days, it is likely that there are no stable orbits for additional planets with orbital periods less than 12.9 days. This leaves two subspaces of the period space where stable orbits could exist, namely, the ones we define above, i.e. orbits between GJ 581 c and d and orbits beyond GJ 581 d.

Because of the division of the period space into two parts, we could draw samples from both subspaces relatively easily. Samplings of the parameter space of a five-Keplerian model revealed a clear probability maximum for the period of the fifth signal at 43.7 days. We plotted the convergence of the period and amplitude parameters of the fifth signal to this solution in Fig. 7 and the posterior probability densities of these two parameters in Fig. 8. This signal was significantly present in the data and the five-Keplerian model was found to have a posterior probability of 460 times that of the four-Keplerian one, which exceeds the detection threshold of 150. We note that around the 2000th member of the Markov chains in Fig. 7 (bottom panel), the variance of the proposal density has been decreased heavily to increase the convergence rate because the adaptation of the proposal density to the posterior is very low when the chain discovers a narrow probability maximum in the parameter space, as happens in this particular case because the width of the posterior density of the period parameter is much narrower than the corresponding parameter space (see Fig. 7, top panel).

Samplings of both five- and six-Keplerian models revealed an additional probability maximum in the period space between 160–200 days with two maxima at roughly 170 and 190 days, but the posterior samplings did not constrain the amplitude of this signal from below, which violates the detection criteria.

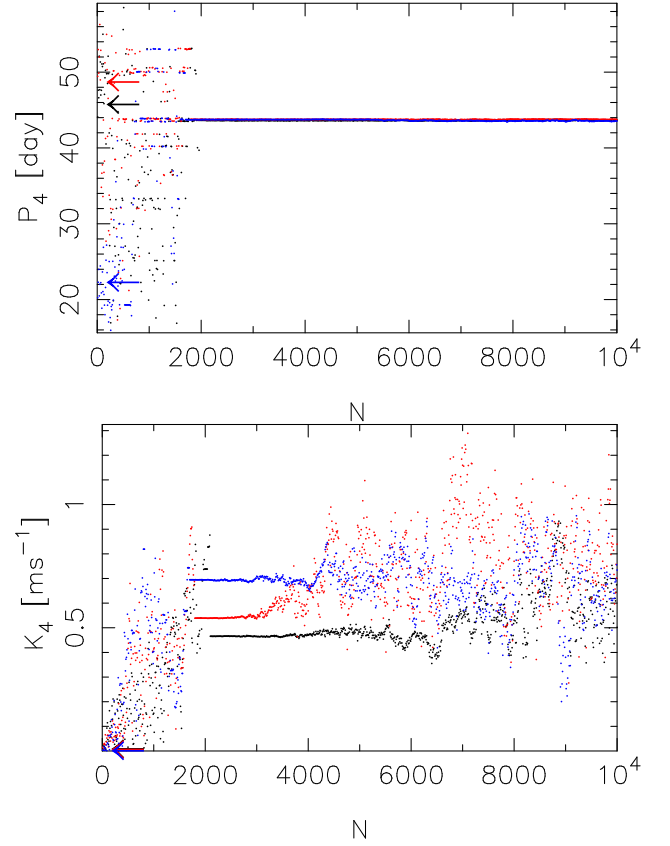


Fig. 7. Convergence of three Markov chains to a signal with a period of 43.7 days and an amplitude of 0.8 ms^{-1} with randomly selected initial states (arrows) in the period space between 12.9 and 66.7 days given the combined HARPS and HIRES data set and a pure white noise model.

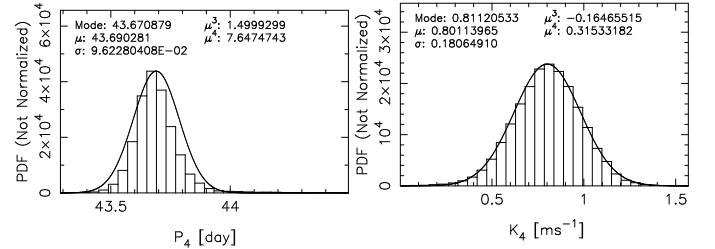


Fig. 8. Posterior probability densities of the period and amplitude of the fifth most significant (fourth shortest periodicity) signal at a period of 43.7 days given the combined HARPS and HIRES data set and a pure white noise model.

Therefore, using the white noise model leads to the conclusion that there are five periodic signals in the combined HARPS and HIRES data. The fact that the 190 day signal observed in the HARPS data alone did not satisfy the detection criteria is unfortunate. While it did appear as a probability maximum given the combined data as well, it could not be constrained as it was for the HARPS data alone. This suggests that this signal is not a genuine Doppler signal of planetary origin but caused by either pure chance in the HARPS velocities or by some (pseudo) periodic source of noise or bias in the HARPS data not accounted for by the pure white noise model.

On the other hand, the nature of the 43.7 day signal in the combined data set is unknown and we cannot rule out its origin as a genuine Doppler signature of planetary origin. It is always

possible that a signal cannot be detected independently in any of the available datasets but only in the combined set because of the better phase-coverage, longer baseline, and a greater number of measurements in the combined set. Yet, it could also be caused by systematic differences between the different data sets, such as small instrumental biases that give rise to weak signals in the combined data set when not accounted for. However, we do not find any indication of such biases and in fact the HARPS and HIRES data sets cannot be shown to contradict with one another given the five-Keplerian model based on the Bayesian model inadequacy criterion of Tuomi et al. (2011). Further, it might be a coincidence that the fifth signal appears at a period that, if of planetary origin, implies a 3:2 commensurability of the corresponding planet with GJ 581 d that has an orbital period of 66.7 days. Such an orbital configuration would likely enable the stability of the system in long term but, while this is compatible with a hypothesis that the 43.7 day signal is caused by a fifth planet candidate in the system, it does not imply the existence of five planets around GJ 581 unless the existence of the fifth signal is supported by future data. Whether this explanation for the 43.7 day signal is valid or not also remains to be investigated with better noise models.

We note that the excess noise levels (estimates of parameter σ_J) of the HARPS and HIRES velocities were 1.50 [1.15, 1.86] and 2.20 [1.63, 2.90] ms^{-1} , respectively, given the best model with $k = 5$.

4.2. Red noise model

We tested the dependence of the observed signals on the selected noise model and the conclusions of Baluev (2012) by taking into account red features in the radial velocity noise using a moving average model (MA). This model was constructed as in Tuomi et al. (2012a) and Tuomi et al. (2012b) but had only one component, making the model a first order MA model with Gaussian white excess noise. Mathematically, this noise model can be written as

$$m_i = r_k(t_i) + \gamma + \epsilon_i + \phi[m_{i-1} - r_k(t_{i-1}) - \gamma] \exp[\alpha(t_{i-1} - t_i)], \quad (3)$$

where each measurement m_i at epoch t_i is described using the superposition of k Keplerian signals (r_k); a reference velocity (γ); Gaussian random variable corresponding to the white noise in the data (ϵ_i) with zero mean and variance $\sigma_i^2 + \sigma_J^2$, where σ_i is the estimated instrument uncertainty from the spectral reduction pipeline and σ_J is an unknown excess noise component sometimes referred to as “stellar jitter”; parameter ϕ is a parameter describing the MA magnitude and the term in brackets corresponds to an exponential decrease of this magnitude as a function of time difference of the i th and $i - 1$ th measurements in accordance with the formulation of Baluev (2012). Parameter α describes the timescale of the correlations in the measurement noise.

As observed by Baluev (2012), the HIRES data alone actually supports the existence of the planets GJ 581 b, c, and e despite the fact that only b and c could be detected using a pure white noise model and Bayesian analyses of the data (Gregory, 2011). Using the MA model described above, the samplings of the parameter space of the three-Keplerian model showed beyond reasonable doubt that these three signals are present in the 122 HIRES velocities of Vogt et al. (2010) by increasing the posterior probability of the model by a factor of 8.2×10^5 greater than that of the two-Keplerian model that corresponded to the two signals at 5.4 and 12.9 days. The third signal with a period

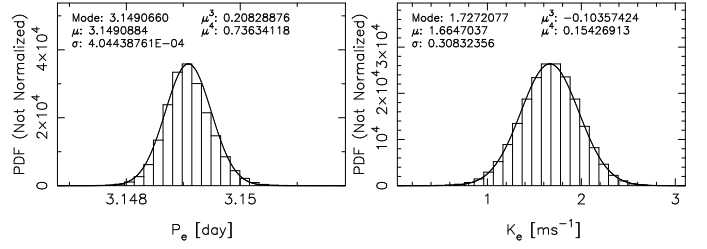


Fig. 9. As in Fig. 1 but for the velocity amplitude and period of GJ 581 e as obtained from the 122 HIRES velocities of Vogt et al. (2010) data using the MA model.

of 3.1 days was constrained well and found consistent with the corresponding signal obtained from the HARPS velocities. We show the posterior densities of the period and velocity amplitude of GJ 581 e in Fig. 9 to demonstrate the significance of this detection.

We could not spot a fourth significant periodic signal in the HIRES data using the posterior samplings and the MA model. However, we performed samplings anyway and spotted three probability maxima for the period of the fourth signal at roughly 70, 300, and 500 days. However, we could not constrain any solutions for the four-Keplerian model and essentially ended up drawing samples from the prior of the period, though the chains were attracted to these three maxima throughout the samplings.

Using the MA model also decreased the estimated amount of white noise in the data. While this white noise was found to be 2.20 [1.63, 2.90] ms^{-1} as estimated with the parameter σ_J in the pure white noise model, it decreased to 1.81 [1.37, 2.72] ms^{-1} for the MA model.

The HARPS velocities, however, did appear to favour a model with five periodic signals. We spotted a 190 day periodicity in the HARPS velocities using the MA model of Eq. (3). Compared to the four-Keplerian model, the five-Keplerian one had 540 times greater posterior probability and the period and amplitude of the signal at 190 days were well constrained in accordance of our detection criteria. Therefore, the HARPS velocities, as was the case for the analyses with the pure white noise model, supports the existence of five independent sources of periodicity corresponding to the four planets reported in Mayor et al. (2009) and one previously unknown signal at roughly 190 day period. We plotted the densities of parameters P and K of this signal in Fig. 10 (top panels) together with the corresponding phase-folded signal with the four planetary signatures of GJ 581 b, c, d, and e, and the MA component of the noise removed (Fig. 10, bottom panel). The white noise parameter σ_J decreased from 1.50 [1.15, 1.86] to 1.19 [0.98, 1.64] ms^{-1} when adding the MA component to the statistical model. This indicates, as was found for the HIRES data, that there are significant correlations in the noise of the HARPS data as well and they have to be accounted for when modelling the velocities.

When analysing the combined HARPS and HIRES data set with the MA model of Eq. (3), we did not have any trouble recovering the signals of the four planets orbiting the star. However, as was the case with the white noise model, we failed to identify the existence of the 190 day signal as a significant one when sampling the period space above 67 days with a five-Keplerian model. We did observe a probability maximum at 190 days in the period space but could not constrain the corresponding signal from above and below in the period space and below in the amplitude space in violation of the detection criteria. Also, we did not find a signal at a period of 43.7 days, which indi-

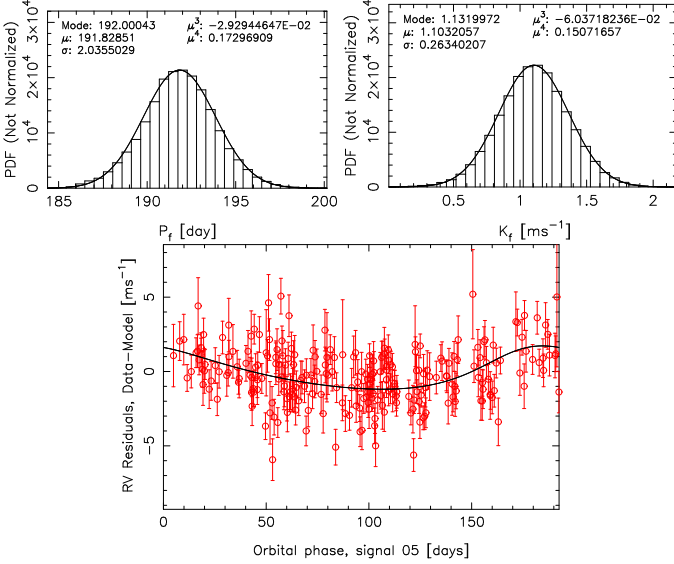


Fig. 10. As in Fig. 3 but for the fifth signal in the Forveille et al. (2011) data given the MA model of Eq. (3).

cates that the corresponding periodicity detected using a pure Gaussian white noise model was in fact most likely a spurious signal generated by insufficient noise modelling. This conclusion is supported by the fact that for $k = 4$ the MA model was found to have a Bayesian evidence of -797.0 on the logarithmic scale, whereas the corresponding Bayesian evidence of the white noise model was -813.1. Further, the log-Bayesian evidence of the white noise model with $k = 5$ was only -806.3, which implies that the MA model with $k = 4$ is the best description of the data out of the set of models tested. We show the parameter MAP estimates and their 99% BCSs of this best model in Table 2.

We note that samplings of a six-Keplerian model did not show any evidence in favour of additional periodic signals in the combined data set.

5. Discussion

Estimating the number of planetary companions orbiting GJ 581, or equivalently, estimating the parameter k favoured by the data for this star, has been attempted by several authors over the past few years (e.g. Bonfils et al., 2005; Udry et al., 2007; Mayor et al., 2009; Vogt et al., 2010; Tuomi, 2011; Gregory, 2011; Vogt et al., 2012; Baluev, 2012). Improvements in statistical analysis techniques have proven efficient means of improving this estimate, i.e. especially showing that some signals have actually been false positives, but the most significant developments have relied on the availability of new high-precision data. We have assessed the number of significant periodic signals in the combined HARPS and HIRES data of GJ 581 and conclude that there is evidence in favour of four signals, corresponding to the four planets orbiting the star observed by Mayor et al. (2009).

By re-analysing sequentially the HARPS data after each observing period, i.e. each full year, we have shown that the Bayesian statistical techniques were at least as sensitive in detecting periodic signals in the data as the more commonly applied periodogram-based analysis tools (Table 1). Moreover, we did not find any significant periodicities in these HARPS data subsets that did not show as significant signals when the amount of data was increased. This implies that the Bayesian detection

criteria (Tuomi, 2012) are not sensitive to false positives arising from noise by chance in practice and any periodic signals that satisfy these criteria are likely genuine signals present in the data.

However, insufficient modelling and possible biases in any single data set can produce false positives that might be interpreted as planetary signals and could contaminate the statistics of exoplanet properties. For instance, the HARPS data was found to favour five periodicities after the sixth full observing period and 197 velocities. This fifth signal at a period of 190 days did not disappear after the inclusion of another 43 velocity measurements and was clearly present in the data published by Forveille et al. (2011). However, another independent data set obtained using the HIRES did not support the existence of this signal. This was evident when the signal could not be detected significantly in the combined HARPS and HIRES data set. Therefore, we conclude that the 190 day periodicity in the HARPS data set, while being significantly present, was likely not of planetary origin but a spurious periodicity caused by some source of (systematic) noise that had variations resembling periodic behaviour. Alternatively, it is possible that the noise model does not describe the HIRES and/or HARPS data well enough, which might result in the inability to detect this signal in the combined set. However, we do not see any evidence of such inadequate modelling as the two data sets are consistent according to the Bayesian model inadequacy criterion of Tuomi et al. (2011).

The combined HARPS and HIRES data set, however, did contain a fifth periodic signal. This time the weak fifth signal occurred at a period of 43.7 days. This result was obtained by modelling the two velocity sets using Gaussian white noise models that are almost a standard way of describing radial velocities in practice. As was shown recently by Baluev (2012), the radial velocity noise of GJ 581 is not white but contains significant correlations (red noise) that, if not accounted for, could mimic periodic behaviour. We found this to be the case and observed that the statistical model containing a first order moving average component improved the model significantly and enabled us to describe the combined data the best with only four periodic signals. Therefore, the 43.7 day signal observed in the combined data set using the white noise model is the most likely a spurious signal caused by insufficient modelling and consequent misinterpretation of the data and is not a genuine Doppler signal of planetary origin. This does not, however, imply beyond reasonable doubt that the 43.7 day periodicity is not a genuine signal because it is possible that the noise model is still not optimal³ and that a better model could enable the recovery of the signal.

These results show that not only can insufficient modelling lead to detections of false positives, but relying on a single data set can also lead to detections of signals that are not supported by other independent measurements. False positives can be caused by these two sources in practice and care is needed to assess whether any given weak signal in any single data set is caused by either of these two effects or not. However, our results suggest that while the Bayesian detection criteria are not particularly sensitive to false positives, they can fall a victim of these two sources of false positives.

Finally, based on our results, it looks like there are four and only four planets orbiting GJ 581. This means that the 32 day signal observed by Vogt et al. (2012) was a false positive, though

³ Actually this is not only possible but rather certain because the origin, nature, time-evolution, and other features of radial velocity noise are very poorly understood at the moment.

Table 2. Four-Keplerian solution of the HARPS and HIRES velocities of GJ 581 together with the parameters quantifying the noise properties. MAP estimates of the parameters and their 99% BCSs.

Parameter	e	b	c	d
P [days]	3.14912 [3.14865, 3.14963]	5.36863 [5.36840, 5.36884]	12.9169 [12.9104, 12.9235]	66.63 [66.31, 66.92]
e	0.16 [0, 0.36]	0.03 [0, 0.05]	0.03 [0, 0.20]	0.17 [0, 0.51]
K [ms ⁻¹]	1.72 [1.36, 2.11]	12.63 [12.24, 13.05]	3.21 [2.64, 3.73]	1.87 [1.12, 2.47]
ω [rad]	0.6 [0, 2π]	0.9 [0, 2π]	1.3 [0, 2π]	6.2 [0, 2π]
M_0 [rad]	5.8 [0, 2π]	5.9 [0, 2π]	1.9 [0, 2π]	0.6 [0, 2π]
a [AU]	0.0285 [0.0266, 0.0302]	0.0407 [0.0380, 0.0431]	0.0731 [0.0684, 0.0773]	0.218 [0.204, 0.231]
$m_p \sin i$ [M _⊕]	1.8 [1.3, 2.3]	15.9 [13.7, 17.9]	5.4 [4.2, 6.5]	5.2 [3.3, 7.5]
Parameter	HARPS	HIRES		
γ [ms ⁻¹]	0.11 [-0.46, 0.61]	0.14 [-0.87, 1.22]		
σ_J [ms ⁻¹]	1.23 [1.06, 1.72]	1.62 [1.15, 2.32]		
α [day ⁻¹]	0.057 [0, 0.446]	0.120 [0, 0.863]		
ϕ	0.49 [0.26, 0.95]	0.91 [0.45, 1]		

the evidence in favour of this signal in their analyses was moderate at most anyway. Also, the periodogram analyses of Baluev (2012) obtained a decreased significance for the companion GJ 581 d because of the inconsistencies in the periodogram analyses described in Section 2.2 and in e.g. Tuomi (2012). However, the HARPS velocities of Forveille et al. (2011) obtained from the spectra using the cross-correlation function method are not optimal (Pepe et al., 2002) but their accuracy can be significantly improved when using more sophisticated spectral reduction methods, such as the TERRA algorithm (Anglada-Escudé & Butler, 2012; Anglada-Escudé & Tuomi, 2012). Improving the noise modelling can also lead to more accurate descriptions of the velocities and therefore enable the detections of one or some of the signals that were only obtained from the HARPS data or with a particular noise model at 190 and 44 day periods. A better understanding of the noise inherent in measuring RVs for M-dwarf stars is essential because current (Barnes et al., 2012) and future (Ramsey et al., 2008; Mahadevan et al., 2009) planet search surveys aim to focus their attention on the inactive and slowly rotating subset of these stars (see Jenkins et al., 2009). For these reasons, and because GJ 581 is still a target of e.g. the HARPS and HIRES exoplanet surveys (according to Vogt et al. (2012) there is already a considerable amount of additional HIRES data), we do not expect that the debate over the number of k for GJ 581 is over. This is especially interesting because the period-space between GJ 581 c and d is known to contain dynamically stable orbits for low-mass planets and happens to correspond to the liquid water habitable zone of the star.

Acknowledgements. M. Tuomi is supported by RoPACS (Rocky Planets Around Cool Stars), a Marie Curie Initial Training Network funded by the European Commission's Seventh Framework Programme. J. S. Jenkins acknowledges funding by Fondecyt through grant 3110004 and partial support from Centro de Astrofísica FONDAP 15010003, the GEMINI-CONICYT FUND and from the Comité Mixto ESO-GOBIERNO DE CHILE.

References

- Anglada-Escudé, G. & Butler, R. P. 2012, *ApJS*, 200, 15A
 Anglada-Escudé, G. & Dawson, R. I. 2011, submitted manuscript (arXiv:1011.0186 [astro-ph.EP])
 Anglada-Escudé, G. & Tuomi, M. 2012, *A&A*, accepted (arXiv:1206.7118 [astro-ph.EP])
 Baluev, R. V., 2011, *Cel. Mech. Dyn. Ast.*, 111, 235
 Baluev, R. V., 2012, *MNRAS*, submitted (arXiv:1209.3154 [astro-ph.EP])
 Barnes, J. R., Jenkins, J. S., Jones, H. R. A., et al. 2012, *MNRAS*, 424, 591
 Berger, J. O. 1980, *Statistical Decision Theory and Bayesian Analysis* (Springer)
 Bonfils, X., Forveille, T., Delfosse, X., et al. 2005, *A&A*, 443, L15
 Cumming, A. 2004, *MNRAS*, 354, 1165
 Forveille, T., Bonfils, X., Delfosse, X., et al., 2011, *ApJ*, submitted (arXiv:1109.2505 [astro-ph.EP])
 Gregory, P. C. 2011, *MNRAS*, 415, 2523
 Haario, H., Saksman, E., & Tamminen, J. 2001, *Bernoulli*, 7, 223
 Hastings, W. 1970, *Biometrika* 57, 97
 Jeffreys, H. 1961, *The Theory of Probability* (The Oxford University Press)
 Jenkins, J. S., Ramsey, L. W., Jones, H. R. A., et al. 2009, *ApJ*, 704, 975
 Kass, R. E. & Raftery, A. E. 1995, *J. Am. Stat. Ass.*, 90, 773
 Lomb, N. R. 1976, *Astrophys. Space Sci.*, 39, 447
 Lovis, C., Ségransan, D., Mayor, M., et al. 2011, *A&A*, 528, A112
 Mahadevan, S., Ramsey, L., Wright, J., et al. 2009, *SPIE*, 7735, 227
 Mayor, M. & Queloz, D. 1995, *Nature*, 378, 355
 Mayor, M., Pepe, F., Queloz, D., et al. 2003, *ESO Messenger*, 114, 20
 Mayor, M., Bonfils, X., Forveille, T., et al. 2009, *A&A*, 507, 487
 Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., et al. 1953, *J. Chem. Phys.*, 21, 1087
 Pepe, F., Mayor, M., Galland, F., et al. 2002, *A&A*, 388, 632
 Ramsey, L. W., Barnes, J., Redman, S. L., et al. 2008, *PASP*, 120, 887
 Scargle, J. D. 1982, *ApJ*, 263, 835
 Tadeu dos Santos, M., Silva, G. G., Ferraz-Mello, S., & Michtchenko, T. A. 2012, *Cel. Mech. Dyn. Ast.*, 113, 49
 Tuomi, M. 2011, *A&A*, 528, L5
 Tuomi, M. 2012, *A&A*, 543, A52
 Tuomi, M. & Jones, H. R. A. 2012, *A&A*, 544, A116
 Tuomi, M. & Kotiranta, S. 2009, *A&A*, 496, L13
 Tuomi, M., Kotiranta, S., & Kaasalainen, M. 2009, *A&A*, 494, 769
 Tuomi, M., Pinfield, D., & Jones, H. R. A. 2011, *A&A*, 532, A116
 Tuomi, M., Jones, H. R. A., Jenkins, J. S., et al. 2012a, *A&A*, submitted
 Tuomi, M., Anglada-Escudé, G., Gerlach, E., et al. 2012b *A&A*, accepted (arXiv:XXXX.XXXX [astro-ph.EP])
 Udry, S., Bonfils, X., Delfosse, X., et al. 2007, *A&A*, 469, L43.
 Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, *Proc. SPIE*, 2198, 362
 Vogt, S., Butler, P., Rivera, E., et al. 2010, *ApJ*, 723, 954.
 Vogt, S. S., Butler, R. P., & Haghighipour, R. N. 2012, *AN*, 333, 1